

The Molecular and Crystal Structure of the Plant Pigment Harunganin

BY RICHARD A. ALDEN* AND GEORGE H. STOUT

Department of Chemistry, University of Washington, Seattle, Washington, U.S.A.

AND JOSEPH KRAUT* AND DARRELL F. HIGH*

Department of Biochemistry, University of Washington, Seattle, Washington, U.S.A.

(Received 21 March 1963)

The molecular and crystal structure of the newly isolated plant pigment harunganin, $C_{30}H_{36}O_4$, has been determined by three-dimensional Patterson superposition and Fourier methods. Refinement was carried out by least-squares computations to a final R index of 9.4% for 3087 reflections with measurable intensities. The unit cell is monoclinic, space group $C2/c$, with $a = 32.90$, $b = 9.30$, $c = 17.80$ Å, and $\beta = 101^\circ 17'$. The locations of 26 out of a total of 36 hydrogen atoms have been determined. The molecule may be described as 7-methyl-4-oxo-2,5,10-trihydroxy-1,1,8-tris(3'-methyl-2'-butenyl)-1,4-dihydroanthracene.

Apparent thermal parameters for the terminal methyl groups of the side chains are unusually high, probably indicating a degree of disorder in the crystal structure.

Introduction

The tree *Harungana madagascariensis* Lam. (Guttiferae) is widely distributed throughout tropical Africa and the Mascarene Islands. When the outer bark is peeled a yellow exudate appears which is sometimes used by the tribesmen for dyeing wood and clothing. An investigation of the pigments contained in this exudate was originally begun by the first two authors as part of a continuing study of the structures of naturally occurring oxygen heterocycles (Stout & Jensen, 1962*b*; Stout, Stout, Welsh & Jensen, 1962). It was thought that the *H. madagascariensis* pigments were likely to be oxygen heterocycles related to those already discovered in other species of the family Guttiferae by Stout & Jensen (1962*a*, *b*) and Stout, Stout, Welsh & Jensen (1962). The compound whose structure is the subject of this paper was one of several pigments isolated from the bark of *H. madagascariensis*. It was given the name harunganin.

An effort was made initially to determine the structure of harunganin by the customary chemical procedures, but in every case attempted derivative-forming or degradative reactions resulted in useless tarry mixtures.

After a protracted but ineffectual investigation along these lines, it was decided to try X-ray diffraction methods. A preliminary short communication describing this work has already appeared (Stout, Alden, Kraut & High, 1962).

Experimental

Samples of the bark obtained from l'Institut National pour l'Etude Agronomique du Congo Belge and from

* Present address: Department of Chemistry, University of California, San Diego, La Jolla, California, U.S.A.

S. B. Penick & Co. were powdered and extracted with light petroleum. The harunganin was separated from a large amount of resinous material by several recrystallizations from methylene chloride-cyclohexane and acetone-cyclohexane. A final slow crystallization from methylene chloride-cyclohexane produced long, bright orange prisms (m.p. 170–171.5 °C) of a size suitable for X-ray study. The crystals were monoclinic, with the unit-cell b axis parallel to the prism axis. Lattice parameters were measured with a precession camera and again on a General Electric XRD-5 single-crystal diffractometer with $Cu K\alpha$ radiation. The precession camera was calibrated against NaCl, a being taken as 5.6387 Å and the mean of $K\alpha_1$ and $K\alpha_2$ as 1.5418 Å. The lattice parameters were found to be

$$a = 32.90 \pm 0.03, \quad b = 9.30 \pm 0.01, \\ c = 17.80 \pm 0.02 \text{ \AA}; \quad \beta = 101^\circ 17' \pm 5'.$$

The uncertainties indicated are estimated standard deviations. A density of 1.150 g.cm⁻³ was measured by flotation in a cyclohexane-carbon tetrachloride mixture. Systematically absent reflections were $h0l$ with l odd and hkl with $h+k$ odd, indicating either space group $C2/c$ or Cc . The molecular weight calculated from these data, assuming eight molecules per unit cell, is 464, which was considered reasonable on the basis of the earlier chemical work. Taken together with the results of an elemental analysis performed on a sample of the crystals, the data thus pointed to a tentative empirical formula of $C_{30}H_{36}O_4$ (mol.wt. 460.6). Obviously, however, the question of whether there was one molecule of this size per asymmetric unit in space group $C2/c$ or two molecules per asymmetric unit in space group Cc remained to be solved.

Copper $K\alpha$ intensity data were collected on the diffractometer by a 2θ -scanning technique, in which the

background correction was made on the basis of the mean background on either side of the reflection peak. The scan time was 20 seconds over a 1.3° interval of 2θ for 2θ values of less than 120° ; for higher 2θ values a longer scan was necessary because of increased α_1 - α_2 splitting. Goniostat settings were calculated on an IBM 650, using U.W. program 0014, for all 5820 possible reflections within the range of the diffractometer ($2\theta \leq 161^\circ$). The data crystal measured approximately $0.08 \times 0.3 \times 0.2$ mm.

Intensity data were reduced with U.W. program 0029 (IBM 650), the usual Lorentz and polarization corrections, but no absorption correction, being applied. At this point the data reduction program also interpolated atomic scattering factors from tables given by Berghuis, Haanappel, Potters, Loopstra, MacGillavry & Veenendaal (1955) for carbon, nitrogen and oxygen, and McWeeny (1951) for hydrogen. The standard deviations of relative intensities were estimated from counting statistics for each reflection. Applying the criterion that $I/\sigma(I)$ must be greater than 2 for a net intensity to be considered significantly different from zero, 3087 or 53% of all reflections scanned were recorded as having measurable intensities. For those reflections not passing the test the data reduction program entered the minimum observable amplitude, 2σ , in place of the observed amplitude.

Determination of the structure

In an attempt to decide between the possible space groups both the Howells, Phillips, & Rogers (1950) and the Wilson (1949) statistical tests were performed (ACA program No. 235, IBM 650). Both indicated the presence of a center of symmetry and therefore favored space group $C2/c$. However, the same calculations performed on data for the acentric steroid androsterone, in space group $P2_1$ (High & Kraut, 1961), also gave a positive test for centrosymmetry, and so the results for harunganin were considered inconclusive. Nevertheless, in its early stages the investigation proceeded on the assumption that $C2/c$ was indeed the correct symmetry, largely because the alternative would have been to attempt the solution of a Cc structure containing 68 light atoms per asymmetric unit. In fact, as will be seen in what follows, it was finally necessary to do just that in order to arrive at the correct structure.

The structure was solved by a combination of machine vector-coincidence methods using a sharpened, origin-removed, three-dimensional Patterson function and straightforward Fourier methods. The Patterson grid size was 0.33 by 0.23 by 0.36 Å. Although a finer grid would have been preferable, the storage capacity of a 32K machine such as the IBM 709, on which these calculations were performed, becomes a limiting factor in problems of this complexity. That the structure contained planar six-

membered rings forming at least a naphthalene-like nucleus was immediately evident upon examination of the Patterson function.

When Patterson superposition methods are applied it is convenient first to locate one or more atoms of the structure with respect to the space group symmetry elements. To do this the self-consistent set of Patterson peaks corresponding to the vectors among these atoms must be found. For two atoms in the asymmetric unit of $C2/c$ the set consists of ten such vectors unrelated by the symmetry of the Patterson space group. At first potential atom-pair locations were sought by visually inspecting the Patterson function for possible vector sets. Later a program (U.W. program SUP 3) was written to carry out the search in a systematic fashion. The program was also capable of searching for any number of vector sets simultaneously.

Four possible atom-pairs were located without great difficulty. The next step was to use these as the basis for a conventional Patterson superposition (U.W. program SUP 2) in which still more of the structure would be recognizable. When this was attempted, however, multiple images of large planar ring systems resulted, but no single structure could be found which did not have unacceptably close intermolecular distances.

After considerable effort had been expended the suspicion arose that possibly the space group was actually Cc rather than $C2/c$. Therefore, it was decided to attempt to find a solution based on superposition functions with just the symmetry elements of Cc despite the difficulty expected in trying to unravel a structure of 68 atoms. Since the same set of symmetry elements is common to $C2/c$ as well, it was hoped to arrive at the correct structure regardless of which turned out to be the true space group. Proceeding on this basis, the SUP 3 program was re-run using now only the Cc symmetry operations, which incidentally give rise to four unique Patterson vectors from an atom-pair instead of ten. An entire plane hexagonal ring was located in this way. When this fragment was now introduced into superposition functions of Cc symmetry calculated with program SUP 2 a large number of further possible atoms were visible in the resulting maps. Of these possible atoms a total of 14 appeared sufficiently reliable to use as the basis of further superposition calculations. Before continuing, however, it was necessary to devise a rapid procedure for distinguishing peaks in the superposition maps which could safely be taken as representing real atoms, since even with an input of fourteen approximately correct atom locations there was still a great deal of interference from multiple images. A fairly useful test was finally devised, based on a comparison of sums of peak heights before and after the introduction of a questionable atom location. Extensive application of this test eventually yielded the positions of 22 additional atoms.

Of the 68 atoms being sought, a total of 36 had now been found, and it was becoming apparent that the asymmetric unit of the Cc space group contained two similar molecules. A three-dimensional Fourier synthesis was next calculated with phases based on these 36 atoms, with the result that 30 additional atoms were immediately evident. The remaining two atoms were revealed by one additional round of Fourier computations.

When a model of the complete 68-atom Cc structure was assembled, it was found to consist of two identical molecules related by a twofold screw axis. The correct space group was thus $C2/c$ after all. The sources of the difficulties experienced in attempting to solve the structure in this space group in the first place have not yet been investigated. It is worthwhile emphasizing, however, that the final success of the combined Patterson superposition-Fourier method when applied to what was, as far as the calculations were concerned, a 68-atom structure, would seem to be a good omen for its ability to handle problems of greater complexity than was previously thought possible.

Refinement

Refinement was carried out entirely by the method of least-squares, the symmetry of the centric space group being used once again. A weighting scheme similar to Hughes's (1941) was employed with the weight assigned to a particular reflection taken as the smaller of the two quantities $[\sigma(F)]^{-2}$ or $(0.05F)^{-2}$, where $\sigma(F)$ was estimated from counting statistics. Unobserved reflections were given zero weight in any round of parameter adjustment in which their calculated amplitudes were less than the minimum observable amplitude inserted by the data reduction program. A single overall scale factor was carried through all refinement computations as an adjustable parameter.

The starting R index, based on atomic positions from the first complete model of the structure, was 37% for 2050 randomly selected observed reflections. Initially, positional parameters and individual isotropic temperature factors were adjusted in alternate cycles by least-squares calculations based on the 827 most intense reflections, the program ORXLS of Busing & Levy (1959) being used. Subsequent cycles were calculated on increasing numbers of randomly selected observed reflections. The scattering factor for carbon was used for all atoms. At the end of this sequence of calculations, the R index stood at 16% on 2374 observed reflections, and the temperature factor B had dropped to about 1 \AA^2 for four of the atoms. It was evident from this together with stereochemical considerations that these four were actually oxygen atoms. The molecular structure of harunganin which had thus emerged is shown schematically in Fig. 1. All subsequent refinement was carried out with the block-diagonal least-squares program of Gantzel, Sparks, & Trueblood (1961) after converting to indi-

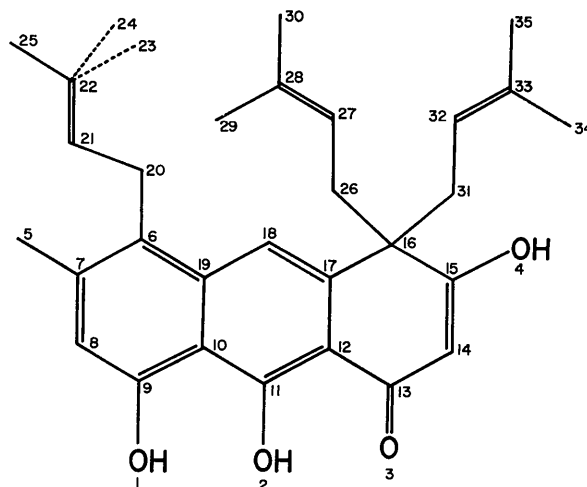


Fig. 1. Schematic structural formula of harunganin molecule.

vidual anisotropic temperature factors, calculating on 2075 observed reflections. As the calculations progressed, the apparent thermal motion of atom C(23) became unreasonably large and anisotropic, with r.m.s. amplitudes along the principal axes of the vibration ellipsoid equal to 0.76 \AA , 0.42 \AA , and 0.28 \AA .

Table 1. Positional parameters of carbon and oxygen atoms, in fractional coordinates $\times 10^4$, and their calculated standard deviations

	x/a	y/b	z/c
O(1)	706 ± 1	4775 ± 4	464 ± 2
O(2)	1395 ± 1	3736 ± 3	1211 ± 2
O(3)	2029 ± 1	2690 ± 3	1996 ± 2
O(4)	2298 ± 1	-2266 ± 3	1945 ± 2
C(5)	-219 ± 2	1763 ± 6	-1294 ± 3
C(6)	428 ± 1	689 ± 5	-476 ± 2
C(7)	177 ± 1	1878 ± 5	-697 ± 3
C(8)	283 ± 1	3246 ± 5	-359 ± 3
C(9)	634 ± 1	3429 ± 5	172 ± 3
C(10)	910 ± 1	2278 ± 4	403 ± 2
C(11)	1288 ± 1	2411 ± 4	930 ± 2
C(12)	1547 ± 1	1239 ± 4	1167 ± 2
C(13)	1925 ± 1	1432 ± 5	1733 ± 2
C(14)	2174 ± 1	230 ± 4	1995 ± 2
C(15)	2075 ± 1	-1094 ± 4	1716 ± 2
C(16)	1705 ± 1	-1450 ± 4	1098 ± 2
C(17)	1434 ± 1	-144 ± 4	854 ± 2
C(18)	1075 ± 1	-304 ± 5	316 ± 2
C(19)	804 ± 1	895 ± 5	75 ± 2
C(20)	306 ± 1	-776 ± 6	-816 ± 3
C(21)	475 ± 2	-1008 ± 6	-1560 ± 3
C(22)	591 ± 2	-2195 ± 7	-1804 ± 3
C(23)	471 ± 6	-3677 ± 14	-1477 ± 8
C(24)	745 ± 6	-3465 ± 14	-1205 ± 10
C(25)	733 ± 3	-2302 ± 10	-2567 ± 4
C(26)	1454 ± 1	-2660 ± 4	1406 ± 3
C(27)	1305 ± 1	-2252 ± 5	2128 ± 3
C(28)	1381 ± 1	-2918 ± 5	2795 ± 3
C(29)	1659 ± 2	-4220 ± 6	2970 ± 3
C(30)	1194 ± 2	-2423 ± 7	3451 ± 4
C(31)	1877 ± 1	-2035 ± 5	394 ± 3
C(32)	2136 ± 2	-979 ± 6	97 ± 3
C(33)	2099 ± 2	-517 ± 6	-642 ± 3
C(34)	1767 ± 2	-1140 ± 10	-1280 ± 3
C(35)	2390 ± 2	556 ± 8	-863 ± 4

Table 2. *Thermal parameters of carbon and oxygen atoms, and their calculated standard deviations*

$$\beta \text{ as given here is defined by: } T = \exp \{ -10^{-5}(\beta_{11}h^2 + \beta_{22}k^2 + \beta_{33}l^2 + \beta_{12}hk + \beta_{13}hl + \beta_{23}kl) \}$$

	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}
O(1)	170 ± 4	1391 ± 48	751 ± 20	252 ± 24	-146 ± 15	-167 ± 49
O(2)	137 ± 4	1051 ± 37	629 ± 15	89 ± 20	-52 ± 12	-98 ± 39
O(3)	121 ± 3	1169 ± 39	511 ± 13	-103 ± 18	-57 ± 11	-195 ± 36
O(4)	124 ± 3	1244 ± 39	506 ± 13	88 ± 19	-49 ± 11	-95 ± 37
C(5)	126 ± 6	2161 ± 86	624 ± 26	190 ± 37	-24 ± 19	235 ± 76
C(6)	90 ± 4	1591 ± 64	435 ± 18	1 ± 27	55 ± 14	74 ± 54
C(7)	99 ± 4	1870 ± 70	412 ± 18	49 ± 29	52 ± 14	125 ± 58
C(8)	116 ± 5	1629 ± 71	553 ± 21	139 ± 30	48 ± 17	38 ± 63
C(9)	125 ± 5	1347 ± 62	482 ± 21	102 ± 28	15 ± 16	9 ± 58
C(10)	108 ± 4	1310 ± 56	418 ± 16	58 ± 26	76 ± 14	-53 ± 50
C(11)	119 ± 5	1147 ± 53	416 ± 17	-58 ± 26	35 ± 15	-47 ± 49
C(12)	93 ± 4	1164 ± 53	384 ± 15	5 ± 23	44 ± 13	-41 ± 46
C(13)	103 ± 4	1391 ± 59	363 ± 15	-62 ± 26	53 ± 13	-97 ± 50
C(14)	103 ± 4	1306 ± 56	386 ± 16	32 ± 25	16 ± 13	161 ± 49
C(15)	97 ± 4	1247 ± 55	395 ± 16	81 ± 25	25 ± 13	126 ± 49
C(16)	109 ± 4	1107 ± 53	407 ± 16	52 ± 25	29 ± 13	-56 ± 48
C(17)	97 ± 4	1181 ± 52	372 ± 16	-14 ± 24	52 ± 13	23 ± 47
C(18)	103 ± 4	1398 ± 60	406 ± 16	-20 ± 26	34 ± 13	-9 ± 52
C(19)	92 ± 4	1441 ± 59	411 ± 16	2 ± 25	74 ± 13	97 ± 50
C(20)	133 ± 6	2045 ± 81	492 ± 21	-51 ± 36	40 ± 17	-233 ± 68
C(21)	184 ± 7	2083 ± 88	515 ± 24	-12 ± 41	117 ± 21	-247 ± 74
C(22)	316 ± 12	2173 ± 106	717 ± 30	139 ± 60	396 ± 32	-303 ± 91
C(23)	570 ± 39	1831 ± 210	802 ± 67	-48 ± 142	745 ± 82	687 ± 188
C(24)	372 ± 31	1378 ± 193	776 ± 108	226 ± 131	115 ± 91	58 ± 237
C(25)	329 ± 15	3644 ± 175	764 ± 36	365 ± 83	364 ± 39	-166 ± 125
C(26)	122 ± 5	1105 ± 54	468 ± 19	-83 ± 27	47 ± 15	54 ± 53
C(27)	106 ± 5	1503 ± 68	626 ± 23	33 ± 28	148 ± 17	337 ± 63
C(28)	151 ± 6	1709 ± 72	609 ± 24	-78 ± 35	239 ± 19	214 ± 68
C(29)	199 ± 8	1862 ± 87	673 ± 28	86 ± 43	122 ± 24	448 ± 79
C(30)	281 ± 11	2399 ± 108	828 ± 35	129 ± 55	512 ± 33	378 ± 98
C(31)	138 ± 5	1454 ± 63	424 ± 18	138 ± 31	109 ± 15	10 ± 55
C(32)	167 ± 6	2055 ± 83	510 ± 21	148 ± 37	218 ± 18	-11 ± 69
C(33)	191 ± 8	2444 ± 97	572 ± 29	451 ± 45	292 ± 24	368 ± 84
C(34)	245 ± 11	4982 ± 203	448 ± 25	-104 ± 79	12 ± 25	931 ± 117
C(35)	260 ± 10	2714 ± 123	887 ± 37	186 ± 59	439 ± 33	487 ± 110

To test the possibility that C(23) was not really present at all, it was removed and two cycles of refinement were run, followed by Fourier and difference-Fourier synthesis calculations. In both cases an obvious broad peak was found where atom C(23) had been. Two possible explanations were evident. On the one hand, C(23) could be disordered in a random way between two loci separated by about 1 Å, with the statistical symmetry retained as $C2/c$. On the other hand, it was possible that the distribution of C(23) atoms was not random among the two loci but rather was systematic in such a way as to destroy the centers of symmetry between molecules. This would make the symmetry strictly only Cc , with two molecules per asymmetric unit, although still very nearly $C2/c$ by virtue of pseudo centers of symmetry which would still be present.

An attempt was made to distinguish between these possibilities by block-diagonal refinement in Cc after perturbing the positions of the C(23) atoms to destroy the centers of symmetry. However, this approach was abandoned when the residual began to increase rather than diminish after each cycle. The program was also modified to include matrix elements containing cross-terms between pseudo-symmetry related atoms, but the same difficulty was encountered.

Table 3. *Positional parameters of hydrogen atoms in fractional coordinates $\times 10^3$*

	x/a	y/b	z/c
H(36)	88 ± 1	465 ± 5	83 ± 3
H(37)	164 ± 1	353 ± 4	155 ± 2
H(38)	253 ± 1	-191 ± 5	237 ± 2
H(39)	-8 ± 2	142 ± 6	-179 ± 3
H(40)	-26 ± 2	263 ± 6	-135 ± 3
H(41)	-39 ± 2	79 ± 6	-106 ± 3
H(42)	7 ± 1	414 ± 5	-45 ± 3
H(43)	242 ± 1	23 ± 5	240 ± 3
H(44)	100 ± 1	-134 ± 4	-1 ± 2
H(45)	-3 ± 1	-85 ± 5	-107 ± 2
H(48)	109 ± 1	-251 ± 5	-246 ± 2
H(50)	57 ± 1	-328 ± 5	-297 ± 3
H(51)	169 ± 1	-348 ± 4	152 ± 2
H(52)	121 ± 1	-282 ± 5	88 ± 2
H(53)	112 ± 2	-139 ± 5	208 ± 3
H(54)	190 ± 2	-394 ± 6	328 ± 3
H(56)	149 ± 2	-499 ± 6	294 ± 3
H(57)	108 ± 2	-139 ± 6	332 ± 3
H(58)	108 ± 2	-327 ± 6	358 ± 3
H(59)	144 ± 2	-227 ± 6	398 ± 3
H(61)	159 ± 1	-220 ± 4	-9 ± 2
H(62)	235 ± 1	-59 ± 5	38 ± 3
H(63)	173 ± 2	-34 ± 6	-175 ± 3
H(65)	262 ± 1	86 ± 5	-34 ± 2
H(66)	212 ± 2	137 ± 6	-123 ± 3
H(67)	257 ± 2	-16 ± 6	-120 ± 3

Table 4. Observed amplitudes and calculated structure factors.

Within each group of constant *k* and *l*, the columns contain from left to right: *h*, 10*F*_o, and 10*F*_c. An asterisk indicates the estimated minimum observable 10*F*_o for reflections which were too weak to be measured

<i>h</i>	10 <i>F</i> _o	10 <i>F</i> _c	<i>h</i>	10 <i>F</i> _o	10 <i>F</i> _c	<i>h</i>	10 <i>F</i> _o	10 <i>F</i> _c
H=0,0			H=0,10			H=0,20		
000			010			020		
001			020			030		
002			030			040		
003			040			050		
004			050			060		
005			060			070		
006			070			080		
007			080			090		
008			090			100		
009			100			110		
010			110			120		
011			120			130		
012			130			140		
013			140			150		
014			150			160		
015			160			170		
016			170			180		
017			180			190		
018			190			200		
019			200			210		
020			210			220		
021			220			230		
022			230			240		
023			240			250		
024			250			260		
025			260			270		
026			270			280		
027			280			290		
028			290			300		
029			300			310		
030			310			320		
031			320			330		
032			330			340		
033			340			350		
034			350			360		
035			360			370		
036			370			380		
037			380			390		
038			390			400		
039			400			410		
040			410			420		
041			420			430		
042			430			440		
043			440			450		
044			450			460		
045			460			470		
046			470			480		
047			480			490		
048			490			500		
049			500			510		
050			510			520		
051			520			530		
052			530			540		
053			540			550		
054			550			560		
055			560			570		
056			570			580		
057			580			590		
058			590			600		
059			600			610		
060			610			620		
061			620			630		
062			630			640		
063			640			650		
064			650			660		
065			660			670		
066			670			680		
067			680			690		
068			690			700		
069			700			710		
070			710			720		
071			720			730		
072			730			740		
073			740			750		
074			750			760		
075			760			770		
076			770			780		
077			780			790		
078			790			800		
079			800			810		
080			810			820		
081			820			830		
082			830			840		
083			840			850		
084			850			860		
085			860			870		
086			870			880		
087			880			890		
088			890			900		
089			900			910		
090			910			920		
091			920			930		
092			930			940		
093			940			950		
094			950			960		
095			960			970		
096			970			980		
097			980			990		
098			990			1000		
099			1000					

Table 4 (cont.)

Table with multiple columns containing numerical data, likely representing atomic coordinates or displacement values. The table is organized into several vertical sections, each with a header label (e.g., H 8+11, H 9+10) and rows of numbers.

Although it was felt that no definite conclusion was warranted on this point, the refinement was now continued based on the assumption of a randomly disordered C(23) in the centrosymmetric space group. The atom was split by assigning two half-weighted carbon atoms, designated C(23) and C(24), to locations separated by 0.76 Å along the direction of greatest thermal motion, and giving them isotropic temperature factors of B=12 Å². The coordinates and anisotropic temperature factors of these were then allowed to refine in the usual way along with the other parameters. The R index was reduced to 12%, while the temperature factors of the half-atoms decreased slightly and their separation increased to about 0.9 Å. During subsequent refinement cycles

these two half-atoms continued to oscillate about their present locations.

A difference-Fourier synthesis was calculated at this stage in order to locate hydrogen atoms. Of a total of 36 in the asymmetric unit, 24 possible hydrogen atoms now could be seen. These had peak heights ranging from 0.37 to 0.82 e.Å⁻³. They were assigned isotropic temperature factors varying from B=2.0 to 5.5 Å², depending on the thermal parameters of the attached atoms, and introduced into the structure. One cycle of block-diagonal refinement was run on the heavy-atom parameters. An additional difference-Fourier synthesis revealed the possible locations of eight of the remaining twelve hydrogen atoms, with peak heights for these ranging from 0.24 to 0.44 e.Å⁻³.

Final adjustment of parameters was made with two cycles of refinement on hydrogen-atom coordinates only, followed by two cycles on heavy-atom coordinates and anisotropic thermal parameters. In this round of calculations, requiring 166 minutes, the mean final adjustment of coordinates in terms of their standard deviations was 1.3σ , and that of thermal parameters was 0.7σ . The maximum final adjustments were 4.5σ in the z coordinate of the disordered atom C(23) and 5.1σ in the β_{33} of side chain atom C(33). The final R index was 9.4% for 3087 reflections with measurable intensities, all taken to have equal weight.

After the last round of refinement calculations, it was concluded that six of the possible hydrogen atoms had unacceptable bond lengths or angles, and so these were omitted from Table 3 and from Fig. 4 since it could not be claimed that they had been located with any degree of assurance. The criterion of acceptability was that bond lengths should not differ from the expected value by more than 0.2 \AA , nor bond angles by more than 12° . These limits represent approximately 2.5 times the calculated standard deviation for hydrogen-atom bond lengths and angles.

Even though the refinement process obviously could have been continued, it was decided that the object of this study had been achieved, the structure of the pigment having been established. Therefore, in

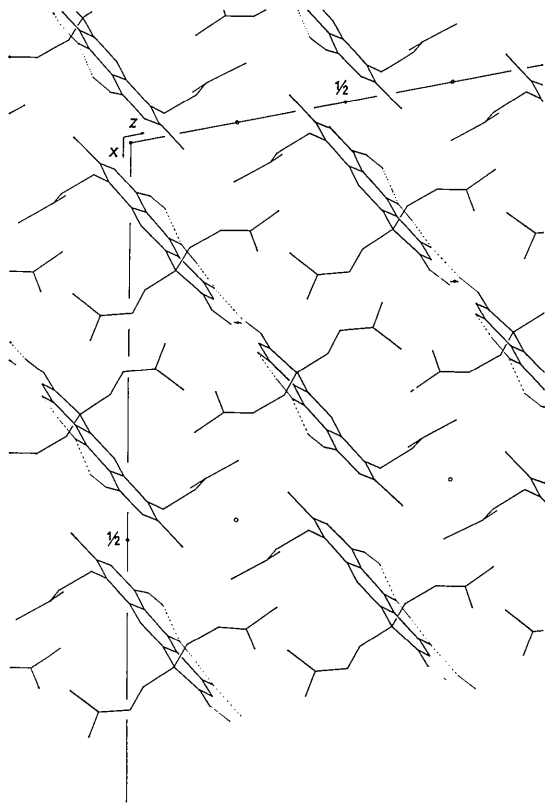


Fig. 2. The crystal structure of harunganin; orthographic projection parallel to the b axis.

view of the large expenditure of computer time required for each cycle of calculations, the refinement was terminated at this point.

Atomic parameters at this stage of refinement are given in Tables 1, 2 and 3. The calculated standard deviations are those put out by the block-diagonal least-squares program. Atom identification numbers are indicated on Fig. 1, 3 and 4 for the heavy atoms, and on Fig. 4 for hydrogen atoms. Observed reflection amplitudes and calculated structure factors are listed in Table 4.

Results and discussion

The crystal structure of harunganin, viewed in orthographic projection along the b axis, is shown in Fig. 2. Three hydrogen bonds, two intramolecular and one between molecules related by the twofold screw axis,

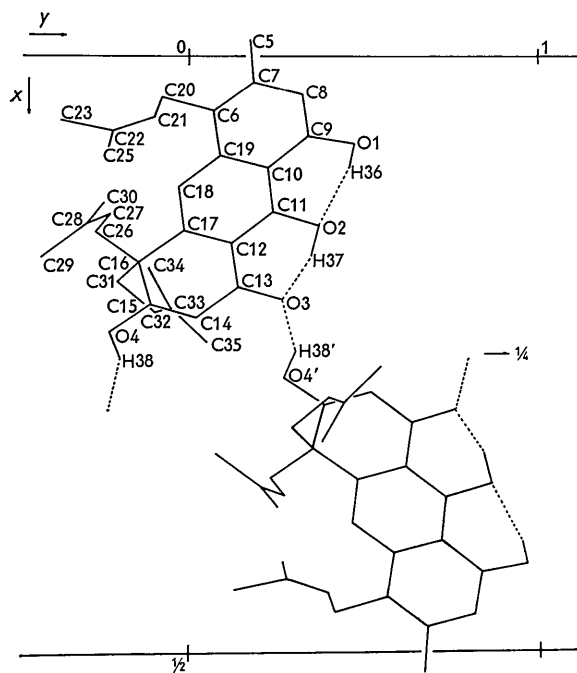


Fig. 3. Orthographic projection parallel to the c axis depicting two molecules situated about a twofold screw axis.

are indicated by dashed lines and the hydrogen atoms involved are also depicted. Other hydrogen atoms are omitted for the sake of clarity. In Fig. 3, just two molecules related by the twofold screw axis are shown in orthographic projection along the c axis. A single complete harunganin molecule, with all hydrogen atoms actually located, is displayed in Fig. 4. This is again an orthographic projection, but parallel to the non-principal direction $[0.4939, 0.5789, 0.0244]$ chosen to give the clearest view of one molecule.

Bond lengths and angles for covalent bonds not involving hydrogen atoms are given in Tables 5 and 6.

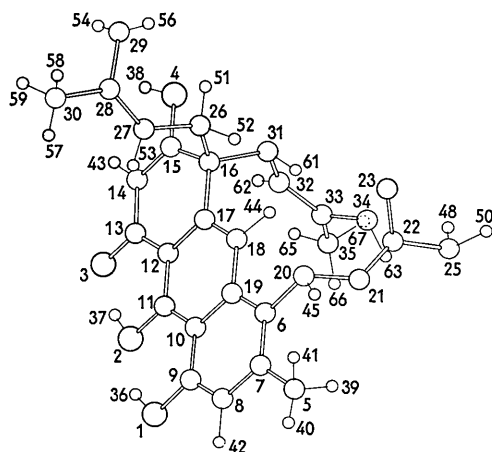


Fig. 4. A single molecule of harunganin showing all hydrogen atoms actually located; orthographic projection parallel to vector direction [0.4939, 0.5789, 0.0244].

Table 5. Bond lengths of covalent bonds not involving hydrogen atoms

Calculated standard deviations are less than 0.01 Å, except for bonds to C(22), where they are less than 0.02 Å. Actual errors may be several times larger in some cases (see text)

Bond	Length
O(1)–C(9)	1.36 Å
O(2)–C(11)	1.35
O(3)–C(13)	1.28
O(4)–C(15)	1.33
C(5)–C(7)	1.52
C(6)–C(7)	1.39
C(6)–C(19)	1.43
C(6)–C(20)	1.51
C(7)–C(8)	1.42
C(8)–C(9)	1.35
C(9)–C(10)	1.41
C(10)–C(11)	1.41
C(10)–C(19)	1.43
C(11)–C(12)	1.40
C(12)–C(13)	1.45
C(12)–C(17)	1.42
C(13)–C(14)	1.41
C(14)–C(15)	1.34
C(15)–C(16)	1.51
C(16)–C(17)	1.52
C(16)–C(26)	1.56
C(16)–C(31)	1.57
C(17)–C(18)	1.37
C(18)–C(19)	1.44
C(20)–C(21)	1.55
C(21)–C(22)	1.27
C(22)–C(23)	1.58
C(22)–C(24)	1.60
C(22)–C(25)	1.52
C(26)–C(27)	1.51
C(27)–C(28)	1.32
C(28)–C(29)	1.51
C(28)–C(30)	1.49
C(31)–C(32)	1.46
C(32)–C(33)	1.37
C(33)–C(34)	1.53
C(33)–C(35)	1.49

Table 6. Bond angles for covalent bonds not involving hydrogen atoms, and their calculated standard deviations

Actual errors may be several times larger in some cases (see text)

Angle	
C(7)–C(6)–C(19)	118.2 ± 0.5°
C(7)–C(6)–C(20)	120.4 ± 0.5
C(19)–C(6)–C(20)	121.4 ± 0.5
C(5)–C(7)–C(6)	121.6 ± 0.6
C(5)–C(7)–C(8)	117.7 ± 0.5
C(6)–C(7)–C(8)	120.7 ± 0.5
C(7)–C(8)–C(9)	120.9 ± 0.6
O(1)–C(9)–C(8)	116.4 ± 0.5
O(1)–C(9)–C(10)	122.2 ± 0.5
C(8)–C(9)–C(10)	121.3 ± 0.5
C(9)–C(10)–C(11)	124.1 ± 0.5
C(9)–C(10)–C(19)	118.2 ± 0.5
C(11)–C(10)–C(19)	117.6 ± 0.5
O(2)–C(11)–C(10)	117.3 ± 0.4
O(2)–C(11)–C(12)	120.1 ± 0.4
C(10)–C(11)–C(12)	122.6 ± 0.5
C(11)–C(12)–C(13)	120.2 ± 0.5
C(11)–C(12)–C(17)	119.4 ± 0.4
C(13)–C(12)–C(17)	120.4 ± 0.4
O(3)–C(13)–C(12)	119.8 ± 0.4
O(3)–C(13)–C(14)	120.5 ± 0.4
C(12)–C(13)–C(14)	119.7 ± 0.5
C(13)–C(14)–C(15)	121.4 ± 0.5
O(4)–C(15)–C(14)	124.0 ± 0.5
O(4)–C(15)–C(16)	111.1 ± 0.4
C(14)–C(15)–C(16)	125.0 ± 0.5
C(15)–C(16)–C(17)	112.1 ± 0.4
C(15)–C(16)–C(26)	108.0 ± 0.4
C(15)–C(16)–C(31)	106.9 ± 0.4
C(17)–C(16)–C(26)	111.0 ± 0.5
C(17)–C(16)–C(31)	109.4 ± 0.4
C(26)–C(16)–C(31)	109.2 ± 0.4
C(12)–C(17)–C(16)	121.2 ± 0.4
C(12)–C(17)–C(18)	119.6 ± 0.5
C(16)–C(17)–C(18)	119.2 ± 0.4
C(17)–C(18)–C(19)	121.4 ± 0.5
C(6)–C(19)–C(10)	120.6 ± 0.5
C(6)–C(19)–C(18)	120.1 ± 0.5
C(10)–C(19)–C(18)	119.3 ± 0.5
C(6)–C(20)–C(21)	110.9 ± 0.5
C(20)–C(21)–C(22)	126.5 ± 0.8
C(21)–C(22)–C(23)	121.3 ± 1.3
C(21)–C(22)–C(24)	119.3 ± 1.2
C(21)–C(22)–C(25)	121.7 ± 0.9
C(23)–C(22)–C(25)	114.4 ± 1.0
C(24)–C(22)–C(25)	115.6 ± 1.1
C(16)–C(26)–C(27)	113.6 ± 0.5
C(26)–C(27)–C(28)	127.7 ± 0.7
C(27)–C(28)–C(29)	124.1 ± 0.7
C(27)–C(28)–C(30)	121.6 ± 0.7
C(29)–C(28)–C(30)	114.3 ± 0.6
C(16)–C(31)–C(32)	112.5 ± 0.5
C(31)–C(32)–C(33)	127.6 ± 0.7
C(32)–C(33)–C(34)	120.7 ± 0.7
C(32)–C(33)–C(35)	121.7 ± 0.8
C(34)–C(33)–C(35)	117.5 ± 0.7

The calculated standard deviations are derived from the standard deviations of coordinates listed in Table 1, with due allowance for error in the unit-cell parameters. Bond lengths and angles are also shown in Fig. 5. It is important to keep in mind that the actual errors in some of these bond lengths and angles (particularly for the side-chains, as discussed below) are probably several times as large as the calculated standard

deviations given in the Tables. This is due partly to the incomplete state of refinement, partly to the tendency for standard deviations obtained from block-diagonal least-squares computations to be somewhat low (Hughes, 1963), and partly to a degree of disorder in the side chains.

The first general observation to be made about the structure of harunganin is that, contrary to expectation, it is *not* an oxygen heterocycle. Although no compound with a closely related structure has been reported previously as far as we know, the molecular skeleton is derivable from that for the natural product emodin (Thomson, 1957).

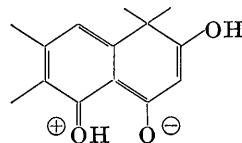
The *gem*-diisopentenyl substitution at C(16) is an unusual feature of the structure, but is not entirely unknown, having been found also in the hop principle lupulon (Verzele & Govaert, 1949).

Bond distances and angles within the naphthalene-like portion of the nucleus are comparable to those

in naphthalene itself (Abrahams, Robertson & White, 1949) and in the similarly substituted oxygen heterocycle rubrofusarin (Stout & Jensen, 1962*b*). The ten atoms comprising the naphthalene nucleus lie very nearly in a plane. The mean distance to the least-squares plane is 0.01 Å, and the maximum is 0.03 Å. The oxygen atoms O(1) and O(2) lie respectively 0.06 Å and -0.08 Å above and below this plane.

As indicated by Fig. 5, bond lengths and angles within the isopentenyl side chains are inconsistent and in poor agreement with the commonly accepted values for unsaturated aliphatic structures. The mean C-C single-bond distance in the side chains is 1.52 Å \pm 0.03 Å for 13 bonds. The mean side-chain tetrahedral angle is 110° \pm 2° for 14 angles, and the mean trigonal angle is 122° \pm 4° for 6 angles. The indicated uncertainties are r.m.s. deviations from the mean among the observed values. These large deviations are probably due to the incomplete state of the refinement, and to thermal motions, free rotations, and disorder in the side chains, for which some allowance has been made, by the use of a split carbon atom, only in the most extreme case of C(23, 24). The apparent thermal parameters for the side-chain atoms in Table 7 are unusually large and anisotropic, particularly at the termini of the isopentenyl chains.

Bond distances to the two phenolic oxygen atoms O(1) and O(2) are 1.36 Å and 1.35 Å, in good agreement with the accepted value of 1.36 Å (Bowen, Donohue, Jenkin, Kennard, Wheatley & Whiffen, 1958). The oxygen functions on the non-aromatic ring, however, have bond distances intermediate between the standard 1.23 Å for ketones and 1.36 Å for enols, namely 1.28 Å for O(3)-C(13) and 1.33 Å for O(4)-C(15). Assignment of the tautomeric form with the hydrogen atom at O(4) is made on the basis of the location of H(38) in the difference-Fourier map, of the possibly significant difference between the two C-O distances, and the similar difference between the two relevant C-C distances. It should be observed that a shift of only 0.7 Å in the location of H(38) would be required to convert the structure to the tautomer with a hydrogen atom on O(3), and it may be asked why there appears to be a preference for the form actually found. The answer may lie in the existence of a non-zero contribution from canonical structures of the type:



Twenty-six of the 36 hydrogen atoms in the molecule have been located with reasonable assurance (the criteria are described under *Refinement*). All the missing hydrogen atoms are in the side chains, as would be expected in view of the high degree of apparent thermal motion exhibited by the side chains. The mean

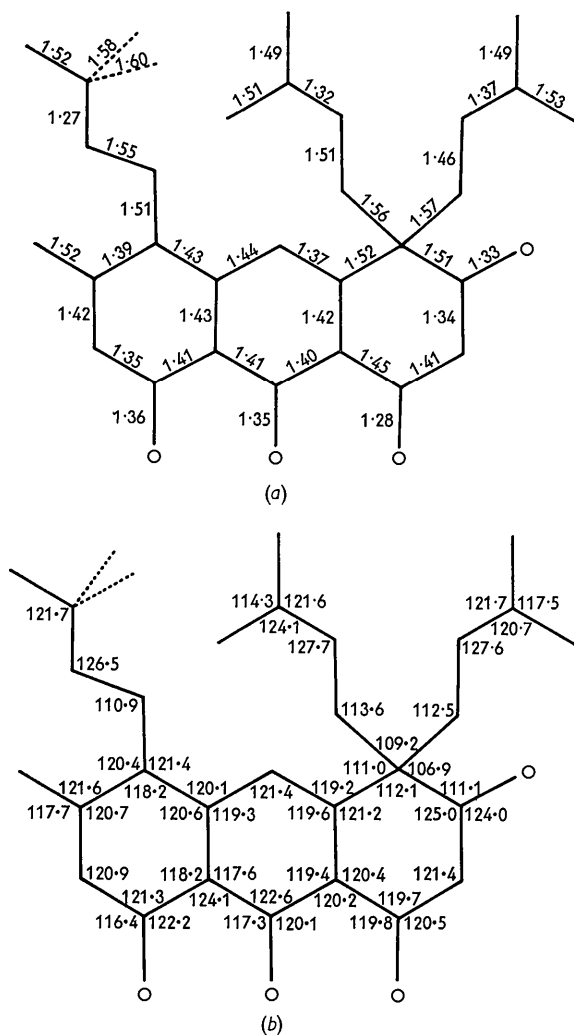


Fig. 5. Bond lengths (a) and angles (b) for the harunganin molecule.

Table 7. *Magnitudes, r.m.s. amplitudes, and direction cosines of the principal axes of thermal vibration ellipsoids*

Axis						Axis							
	i	B_i	U_i	c_{ia}	c_{ib}	c_{ic}		i	B_i	U_i	c_{ia}	c_{ib}	c_{ic}
O(1)	1	12.60	0.40	-0.6730	-0.1928	0.8320	C(19)	1	5.32	0.26	-0.0942	0.6897	0.7225
	2	5.81	0.27	-0.5359	-0.5383	-0.5330		2	4.69	0.24	0.0743	0.7240	-0.6870
	3	3.96	0.22	-0.5098	-0.8204	0.1542		3	3.83	0.22	-0.9928	-0.0113	0.0771
O(2)	1	9.35	0.34	-0.6057	-0.1084	0.8916	C(20)	1	7.50	0.31	-0.1368	-0.8565	0.5149
	2	4.92	0.25	-0.7617	-0.2074	-0.4529		2	6.44	0.29	0.7451	-0.4200	-0.6538
	3	3.51	0.21	-0.2301	0.9722	0.0028		3	4.97	0.25	0.6527	0.3000	0.5545
O(3)	1	7.92	0.32	-0.6566	-0.0372	0.8672	C(21)	1	7.71	0.31	-0.7370	-0.6092	0.4313
	2	4.87	0.25	-0.5572	0.6971	-0.3335		2	7.65	0.31	-0.6674	0.6230	-0.2694
	3	3.27	0.20	0.5084	0.7160	0.3697		3	5.80	0.27	-0.0665	0.4961	0.8620
O(4)	1	7.91	0.32	-0.6718	-0.1740	0.8375	C(22)	1	13.79	0.42	0.8782	0.0726	0.2918
	2	4.43	0.24	-0.4313	-0.6984	-0.4758		2	8.71	0.33	0.2677	0.7445	-0.6522
	3	3.96	0.22	0.6022	-0.6943	0.2687		3	5.96	0.27	-0.3963	0.6637	0.6996
C(5)	1	8.75	0.33	-0.4926	0.1250	0.9410	C(23)	1	24.61	0.56	0.9128	0.0129	0.2218
	2	7.96	0.32	-0.3749	-0.9234	-0.0072		2	8.94	0.34	0.3114	-0.6688	-0.7230
	3	4.25	0.23	0.7850	-0.3635	0.3385		3	4.23	0.23	-0.2644	-0.7434	0.6543
C(6)	1	5.71	0.27	-0.1596	0.7675	0.6401	C(24)	1	16.15	0.45	-0.9901	-0.1156	0.2720
	2	5.22	0.26	0.2138	0.6408	-0.7649		2	9.43	0.35	-0.0731	-0.0612	-0.9619
	3	3.74	0.22	0.9638	-0.0151	0.0726		3	4.60	0.24	0.1200	-0.9914	0.0276
C(7)	1	6.59	0.29	0.0521	0.9701	0.2223	C(25)	1	15.46	0.44	0.7706	0.5904	0.0845
	2	5.13	0.25	0.4469	0.1896	-0.9448		2	11.76	0.39	-0.3976	0.7397	-0.4544
	3	4.03	0.23	0.8931	-0.1515	0.2407		3	7.66	0.31	-0.4979	0.3227	0.8868
C(8)	1	7.09	0.30	-0.4167	-0.1613	0.9589	C(26)	1	6.21	0.28	-0.6081	0.1942	0.8738
	2	6.09	0.28	-0.4100	-0.8450	-0.2565		2	4.95	0.25	-0.7278	0.2509	-0.4835
	3	4.26	0.23	0.8113	-0.5099	0.1217		3	3.65	0.21	-0.3171	-0.9483	0.0511
C(9)	1	6.82	0.29	-0.6740	-0.1841	0.8335	C(27)	1	8.06	0.32	0.0423	0.3582	0.9064
	2	5.14	0.26	-0.4645	-0.6475	-0.5016		2	4.80	0.25	-0.1380	0.9267	-0.3159
	3	4.17	0.23	0.5745	-0.7395	0.2317		3	4.20	0.23	-0.9895	-0.1139	0.2805
C(10)	1	5.24	0.26	-0.3803	-0.4139	0.8855	C(28)	1	8.15	0.32	0.4252	0.1595	0.7905
	2	4.77	0.25	-0.5886	-0.5813	-0.4357		2	6.42	0.29	0.5458	-0.8304	-0.2165
	3	4.14	0.23	-0.7133	0.7005	0.1616		3	4.51	0.24	-0.7220	-0.5339	0.5729
C(11)	1	5.74	0.27	-0.7493	0.0798	0.7913	C(29)	1	9.03	0.34	-0.1376	0.4767	0.8784
	2	4.67	0.24	-0.5721	0.4220	-0.5778		2	8.40	0.33	0.9839	0.1665	-0.1295
	3	3.81	0.22	0.3335	0.9031	0.2001		3	5.59	0.27	0.1140	-0.8631	0.4602
C(12)	1	4.91	0.25	-0.4282	-0.1629	0.9555	C(30)	1	14.28	0.43	0.6884	0.2058	0.5474
	2	4.01	0.23	0.0980	-0.9862	-0.1502		2	8.17	0.32	0.3165	-0.9480	-0.0940
	3	3.80	0.22	0.8984	0.0298	0.2540		3	6.15	0.28	-0.6526	-0.2426	0.8316
C(13)	1	5.04	0.25	0.3110	-0.8950	0.2528	C(31)	1	6.29	0.28	0.8164	0.5497	0.0137
	2	4.69	0.24	0.7036	-0.0096	-0.8345		2	5.14	0.26	0.0647	0.2176	-0.9677
	3	3.92	0.22	0.6389	0.4460	0.4896		3	4.45	0.24	-0.5739	0.8065	0.2517
C(14)	1	5.52	0.26	-0.5723	0.3398	0.8439	C(32)	1	8.12	0.32	0.6715	0.5998	0.2954
	2	4.63	0.24	0.5740	0.8159	-0.0450		2	6.81	0.29	-0.2697	0.7448	-0.5458
	3	3.67	0.22	0.5856	-0.4677	0.5346		3	4.88	0.25	-0.6902	0.2924	0.7841
C(15)	1	5.28	0.26	-0.5071	0.1441	0.9326	C(33)	1	11.49	0.38	0.5969	0.6720	0.3131
	2	4.68	0.24	-0.5390	-0.8224	-0.0733		2	6.40	0.28	-0.1668	0.6383	-0.7043
	3	3.45	0.21	0.6725	-0.5505	0.3535		3	4.76	0.25	-0.7848	0.3754	0.6371
C(16)	1	5.60	0.27	-0.6581	-0.2078	0.8385	C(34)	1	18.26	0.48	-0.1885	0.9525	0.2714
	2	4.25	0.23	-0.6712	-0.2736	-0.5443		2	10.57	0.37	0.9763	0.2082	-0.2495
	3	3.71	0.22	-0.3412	0.9391	0.0268		3	4.67	0.24	0.1066	-0.2222	0.9295
C(17)	1	4.71	0.24	-0.4772	0.1786	0.9372	C(35)	1	13.87	0.42	0.5764	0.3646	0.6044
	2	4.09	0.23	-0.2531	0.9096	-0.2736		2	8.80	0.33	-0.4046	0.9049	-0.0506
	3	3.94	0.22	0.8414	0.3756	0.2163		3	7.23	0.30	-0.7100	-0.2195	0.7951
C(18)	1	5.36	0.26	-0.5775	0.0862	0.9091							
	2	4.86	0.25	-0.0700	0.9854	-0.1384							
	3	4.07	0.23	0.8133	0.1466	0.3930							

C-H distance for the hydrogen atoms actually located is 1.07 Å with an r.m.s. deviation from the mean of 0.12 Å. Definite location of hydrogen atoms attached to ring atoms C(8), C(14) and C(18) lends further support to the chemical structure of harunganin presented here.

In addition to the intermolecular hydrogen bond between O(4) of one molecule and O(3) of the next, there are also intramolecular hydrogen bonds between O(1) and O(2), and between O(2) and O(3). The O-O distances in these hydrogen bonds are 2.61 Å, 2.58 Å and 2.47 Å, all with calculated standard deviations of less than 0.01 Å. Although the last is unusually short, a similarly short O-O hydrogen bond distance of 2.45 Å has been found by Downie & Speakman (1954) in ammonium hydrogen disalicylate monohydrate.

The magnitudes, r.m.s. amplitudes, and direction cosines of the principal axes of thermal vibration ellipsoids are given in Table 7. The quantities c_{ia} , c_{ib} , and c_{ic} are cosines of the angles between the i th principal ellipsoid axis and the a , b , and c axes of the unit cell. The B_i are given in Å² and U_i in Å. All atoms were found to be vibrating with significant anisotropy using the criterion that if at least one of the β_{ij} for the atom in question should differ by more than 2σ from the value it would have if the atom were actually vibrating isotropically with B equal to the mean principal axis B_i , it was considered significantly anisotropic.

In very approximate terms, the apparent splitting of C(23, 24) can be ascribed to an oscillatory rotation about the bond C(20)-C(21). This can be seen from the relatively modest amplitudes of thermal motion given for C(20) and C(21) in Table 7, the larger amplitude of motion in the x direction shown for C(22) and C(25), and the fact that the splitting and thermal motion of C(23, 24) is essentially in the x direction.

This investigation was supported by research grants RG 4859 and CY 4082 from the U.S. Public Health

Service, and by the grants of free computer time from the Research Computer Laboratory at the University of Washington.

References

- ABRAHAMS, S. C., ROBERTSON, J. M. & WHITE, J. G. (1949). *Acta Cryst.* **2**, 233, 238.
- BERGHUIS, J., HAANAPPEL, IJ. M., POTTERS, M., LOOPSTRA, B. O., MACGILLAVRY, C. H. & VEENENDAAL, A. L. (1955). *Acta Cryst.* **8**, 478.
- BOWEN, H. J. M., DONOHUE, J., JENKIN, D. G., KENNARD, O., WHEATLEY, P. J. & WHIFFEN, D. H. (1958). *Tables of Interatomic Distances and Configuration in Molecules and Ions*. Special Publication No. 11. London: The Chemical Society.
- BUSING, W. R. & LEVY, H. A. (1959). *A Crystallographic Least Squares Refinement Program for the IBM-704*. U.S. Atomic Energy Commission Publication ORNL 59-4-37.
- DOWNIE, T. C. & SPEAKMAN, J. C. (1954). *J. Chem. Soc.* p. 787.
- GANTZEL, P. K., SPARKS, R. A. & TRUEBLOOD, K. N. (1961). Private communication.
- HIGH, D. F. & KRAUT, J. (1961). *Program and Abstracts, A.C.A. Annual Meeting*, p. 47.
- HOWELLS, E. R., PHILLIPS, D. C. & ROGERS, D. (1950). *Acta Cryst.* **3**, 210.
- HUGHES, E. W. (1941). *J. Amer. Chem. Soc.* **63**, 1737.
- HUGHES, E. W. (1963). Private communication.
- MCWEENY, R. (1951). *Acta Cryst.* **4**, 513.
- STOUT, G. H. & JENSEN, L. H. (1962a). Private communication.
- STOUT, G. H. & JENSEN, L. H. (1962b). *Acta Cryst.* **15**, 451.
- STOUT, G. H., STOUT, V. F., WELSH, M. J. & JENSEN, L. H. (1962). *Tetrahedron Letters*, No. 13, 541.
- STOUT, G. H., ALDEN, R. A., KRAUT, J. & HIGH, D. F. (1962). *J. Amer. Chem. Soc.* **84**, 2653.
- THOMSON, R. H. (1957). *Naturally Occurring Quinones*. p. 191. London: Butterworth Scientific Publications.
- VEREZELE, M. & GOVAERT, F. (1949). *Bull. soc. chim. Belg.* **58**, 432.
- WILSON, A. J. C. (1949). *Acta Cryst.* **2**, 318.